Using chemometrics in assessing langat river water quality and designing a cost-effective water sampling strategy.

ABSTRACT

Seasonally dependent water quality data of Langat River was investigated during the period of December 2001 – May 2002, when twenty-four monthly samples were collected from four different plots containing up to 17 stations. For each sample, sixteen physico-chemical parameters were measured in situ. Multivariate treatments using cluster analysis, principal component analysis and factorial design were employed, in which the data were characterised as a function of season and sampling site, thus enabling significant discriminating factors to be discovered. Cluster analysis study based on data which were characterised as a function of sampling sites showed that at a chord distance of 75.25 two clusters are formed. Cluster I consists of 6 samples while Cluster II consists of 18 samples. The sampling plots from which these samples were taken are readily identified and the two clusters are discussed in terms of data variability. In addition, varimax rotations of principal components, which result in varimax factors, were used in interpreting the sources of pollution within the area. The work demonstrates the importance of historical data, if they are available, in planning sampling strategies to achieve desired research objectives, as well as to highlight the possibility of determining the optimum number of sampling stations which in turn would reduce cost and time of sampling.

Keyword: Chemometrics; Cluster analysis; Factorial design; Principal component analysis.